UNIVERSITÉ DE MONTRÉAL

NONLINEAR SPATIAL PRICE EQUILIBRIUM

by

Jacques A. Ferland, Lourdes Zubieta,
and Heinz Spiess

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

NONLINEAR SPATIAL PRICE EQUILIBRIUM

by

Jacques A. Ferland, Lourdes Zubieta,
and Heinz Spiess

This research was supported by the Natural Sciences and Engineering Research Council of Canada (N.S.E.R.C.) (grant A8312) and F.C.A.R. of the Gouvernement du Québec (grant 0899).

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche.

Université de Montréal

Centre de recherche sur les transports - Publication #477

Département d'informatique et de recherche opérationnelle - Publication #603

June 1986

ABSTRACT

The nonlinear spatial price equilibrium problem for a single commodity is solved with Gauss-Seidel type techniques. Linear approximations of the equilibrium conditions (or equivalently, quadratic approximations of the associated optimization problem) are used to evaluate variables sequentially either one at a time, or by blocks of flow variables out of common supply (demand) node. These methods are of the Gauss-Seidel-Newton type. Numerical results are given to compare these with MINOS, an augmented Lagrangian technique.

- - - - - - - - - - - - -

RÉSUMÉ

La problème d'équilibre spatial avec un seul produit est traité avec des méthodes de type Gauss-Seidel. Ainsi, nous utilisons des approximations linéaires des conditions d'équilibre (ou de façon équivalente, des approximations quadratiques du problème d'optimisation associé) pour évaluer séquentiellement les variables, soit une à la fois, soit par groupe de variables rattachées à un même point d'approvisionnement (de demande). Ces méthodes se classent donc comme de type Gauss-Seidel-Newton. Des résultats numériques sont présentés pour les comparer avec le logiciel MINOS qui est une méthode de lagrangien augmenté.

1. INTRODUCTION

The spatial price equilibrium problem for a single commodity is characterized by a set of supply nodes i, $1 \leq i \leq m$, and a set of demand nodes j, $1 \leq j \leq n$, for a single commodity. For each supply node i ($1 \leq i \leq m$), d_i denotes the quantity supplied and $S_i(d_i)$ the supply function specifying the price of the commodity at node i. For each demand node j ($1 \leq j \leq n$), b_j denotes the quantity demanded and $I_j(b_j)$ the inverse demand function specifying the cost of the commodity at node j. Furthermore, the commodity is transported from supply nodes to demand nodes; t_{ij} denotes the quantity transported from i to j and $c_{ij}(t_{ij})$ is the transportation cost from i to j.

The equilibrium conditions are as follows :

$$S_i(d_i) + c_{ij}(t_{ij}) - I_j(b_j) \geq 0 \qquad 1 \leq i \leq m,\ 1 \leq j \leq n \qquad (1.1)$$

$$t_{ij}\,(S_i(d_i) + c_{ij}(t_{ij}) - I_j(b_j)) = 0 \qquad 1 \leq i \leq m,\ 1 \leq j \leq n \qquad (1.2)$$

$$d_i = \sum_{j=1}^{n} t_{ij} \qquad 1 \leq i \leq m \qquad (1.3)$$

$$b_j = \sum_{i=1}^{m} t_{ij} \qquad 1 \leq j \leq n \qquad (1.4)$$

$$t_{ij} \geq 0 \qquad 1 \leq i \leq m,\ 1 \leq j \leq n \ . \qquad (1.5)$$

Conditions (1.1) and (1.2) indicate that the flow of commodity t_{ij} between i and j can be positive only if the price at the supply node added to the transportation cost is equal to the price at the demand node. Otherwise, when the sum of the former prices exceed the later, the flow remains at zero.

Spatial price equilibrium problems have retained the attention of several authors since the earlier works of Samuelson (1952) and Takayama and Judge (1964, 1970, 1971). The approaches proposed in the literature include mathematical programming, complementarity pivot theory and variational inequalities.

Within the mathematical programming approach, a fixed point method is proposed by McKinnon (1975), whereas Polito et al. (1980) suggest Benders' decomposition. Florian and Los (1982) proposed the adaptation of nonlinear programming techniques and more recently, Nagurney presented a solution schema based on Gauss-Seidel decomposition (1985).

Other authors prefer to solve the equilibrium conditions using complementary pivot theory. Mathiesen (1974), Asmuth et al. (1979), Pang and Lee (1981) and Jones et al. (1985) analyze the problem in which the functions are linear. Then Friesz et al. (1983) and Jones et al. (1985) extend the approach to the nonlinear version of the problem using successive linearization techniques.

For the multicommodity problems Florian and Los (1982) reformulated the equilibrium conditions as a variational inequality problem, Friez et al. (1981) proposed a Jacobi diagonalization, and Pang (1983) suggested a linearized Gauss-Seidel decomposition method.

Now assume the following hypotheses :

i) $S_i(d_i)$ is continuous, bounded, and monotone increasing, $1 \leq i \leq m$;

ii) $I_j(b_j)$ is continuous, bounded, and monotone decreasing, $1 \leq j \leq n$;

iii) $c_{ij}(t_{ij})$ is continuous, bounded, and monotone increasing, $1 \leq i \leq m$, $1 \leq j \leq n$.

It is easy to verify then that the system (1.1) to (1.5) represents the Kuhn-Tucker conditions (necessary and sufficient) for the problem

$$\text{Min} \sum_{i=1}^{m} \int_0^{d_i} S_i(x)dx + \sum_{i=1}^{m}\sum_{j=1}^{n} \int_0^{t_{ij}} c_{ij}(x)dx - \sum_{j=1}^{n} \int_0^{b_j} I_j(y)dy$$

$$\text{Subject to} \quad \sum_{j=1}^{n} t_{ij} = d_i \qquad 1 \leq i \leq m$$

$$\sum_{i=1}^{m} t_{ij} = b_j \qquad 1 \leq j \leq n \qquad (1.6)$$

$$t_{ij} \geq 0 \qquad 1 \leq i \leq m,\ 1 \leq j \leq n.$$

Equivalently, this problem can be specified in terms of variables t_{ij} uniquely as follows

$$\text{Min } g(t)=\sum_{i=1}^{m}\int_{0}^{\sum_{j=1}^{n} t_{ij}} S_i(x)dx+\sum_{i=1}^{m}\sum_{j=1}^{n}\int_{0}^{t_{ij}} c_{ij}(x)dx-\sum_{j=1}^{n}\int_{0}^{\sum_{i=1}^{m} t_{ij}} I_j(y)dy \qquad (1.7)$$

$$\text{s.t.} \qquad t_{ij} \geq 0 \qquad 1\leq i\leq m,\ 1\leq j\leq n.$$

This problem has a simple feasible domain reducing to the non-negative orthant.

Standard nonlinear programming approaches can be used to solve problem (1.7). In fact, several variants of the Frank-Wolfe technique, the projected gradient, the reduced gradient, the simplex convex technique, and a quadratic approximation technique were implemented for this problem. They were compared with MINOS, an augmented lagrangian technique. This last technique was by far the most efficient (see Zubieta, 1986).

In this paper we introduce Gauss-Seidel type techniques to solve (1.7). The basic idea is to use quadratic approximations of (1.7)(or equivalently, linear approximations of the equilibrium conditions (1.1) to (1.5)) to evaluate variables sequentially either one at a time, or by blocks. Appropriate conditions have to be verified to ensure the convergence of these techniques.

In Section 2, we analyze the variant in which the variables are modified one at a time. The algorithm embeds the Newton approach into the Gauss-Seidel technique for each variable separately as the (i,j) relation (1.2) linearized with respect to t_{ij} is used to evaluate this variable. This is to be contrasted with the standard Gauss-Seidel-Newton approach (see Ortega and Rheinboldt, 1970) where the (i,j) relation of the linearized system is used to evaluate t_{ij}. Block versions where all flow variables from a common supply node are modified simultaneously are then analyzed in Section 3. Several variants are studied in the context of active

sets methods (see Luenberger, 1984). Finally, in Section 4, numerical results are given to compare these methods with MINOS (see Murtagh and Saunders, 1978). These demonstrate the efficiency of two of the block versions.

2. O-D PAIRS DECOMPOSITION

In this algorithm a Gauss-Seidel-Newton type approach is used to solve the equilibrium conditions where the variables t_{ij} are modified one at a time.

Since the algorithm does not depend on the specific structure of the spatial equilibrium problem, it will be defined for the general problem

$$\underset{x>0}{\text{Min}}\ f(x) \tag{2.1}$$

where $f : R^n \to R$, $f \in C^2$, bounded from below and convex. The algorithm is an iterative procedure for generating a solution to the Kuhn-Tucker optimality conditions :

$$\begin{aligned} \nabla f(x) &\geq 0 \\ x^T \nabla f(x) &= 0 \\ x &\geq 0 \end{aligned} \tag{2.2}$$

Given $x^k = [x_1^k, x_2^k, \ldots, x_n^k]^T > 0$ (feasible for (2.1)), x^{k+1} is generated as follows : for $1<i<n$

i) determine a linear approximation of the i^{th} relation in (2.2)

$$\nabla_i f(x^{k,i}) + (x_i - x_i^k)\, \nabla_{ii} f(x^{k,i})$$

where $x^{k,i} = [x_1^{k+1}, x_2^{k+1}, \ldots, x_{i-1}^{k+1}, x_i^k, \ldots, x_n^k]^T$, $\nabla_i f(x^{k,i})$ and $\nabla_{ii} f(x^{k,i})$ are the first and second order partial derivative of f at $x^{k,i}$ with respect to x_i, respectively;

ii) assuming that $\nabla_{ii} f(x^{k,i}) > 0$, let

$$x_i^{k+1} = \text{Max}\left\{0,\ x_i^k - \frac{\nabla_i f(x^{k,i})}{\nabla_{ii} f(x^{k,i})}\right\} . \tag{2.3}$$

The procedure is initiated with some $x^0 > 0$ and terminates when $x^{k+1}=x^k$.

The algorithm introduces the Newton approach into the Gauss-Siedel technique since x_i^{k+1} specified in (2.3) is an optimal solution of the following one-dimensional problem :

$$\underset{x_i>0}{\text{Min}} \quad \nabla_i f(x^{k,i})\,(x_i-x_i^k) + \tfrac{1}{2}\,(x_i-x_i^k)^2\,\nabla_{ii} f(x^{k,i}) \quad . \qquad (2.4)$$

The convergence proof will rely heavily on this observation. Indeed, since x_i^{k+1} is an optimal solution of (2.4) it follows that for all $y_i > 0$

$$[\nabla_i f(x^{k,i}) + (x_i^{k+1}-x_i^k)\,\nabla_{ii} f(x^{k,i})]\,(x_i^{k+1}-y_i) \leq 0 \qquad (2.5)$$

and in particular, for $y_i = x_i^k$

$$\nabla_i f(x^{k,i})(x_i^{k+1}-x_i^k) \leq -(x_i^{k+1}-x_i^k)^2\,\nabla_{ii} f_i(x^{k,i}) \leq 0 \qquad (2.6)$$

since $\nabla_{ii} f_i(x^{k,i}) > 0$.

To analyze global convergence, the following hypotheses are required.

H1. $f : R^n \to R$, $f \in C^2$, bounded from below and convex.

H2. $\nabla_{ii} f(x) > 0$ for all $x>0$, $1 \leq i \leq n$.

H3. (Lipschitz condition) There exists a scalar $L > 0$ such that for all $x,y>0$

$$\|\nabla f(x) - \nabla f(y)\| \leq L\,\|x-y\| \quad ,$$

and a scalar $\varepsilon > 0$ such that for all $x > 0$, and $1 \leq i \leq n$

$$0 \leq L/2 + \varepsilon < \nabla_{ii} f(x) \quad .$$

The first result indicates the monotonicity of the objective function throughout the procedure.

Proposition 2.1. Assume that assumptions H1, H2 and H3 hold.

i) For all $k \geq 1$, $1 \leq i \leq n$,

$$f(x^{k,i+1}) \leq f(x^{k,i}) \tag{2.7}$$

with the strict inequality holding unless $x^{k,i+1} = x^{k,i}$. (Note that by convention $x^{k,n+1} = x^{k+1,1}$.)

ii) For all $k \geq 1$, $1 \leq i \leq n$

$$\lim_{k \to \infty} | x_i^{k+1} - x_i^k | = 0 \tag{2.8}$$

Proof.

i) From the Mean Value Theorem

$$f(x^{k,i+1}) = f(x^{k,i}) + \nabla f(x^{k,i})^T (x^{k,i+1} - x^{k,i}) + \int_0^1 \{\nabla f[x^{k,i} + t(x^{k,i+1} - x^{k,i})] - \nabla f(x^{k,i})\}^T (x^{k,i+1} - x^{k,i}) \, dt .$$

Hence

$$f(x^{k,i}) - f(x^{k,i+1}) \geq -\nabla_i f(x^{k,i}) (x_i^{k+1} - x_i^k) - \int_0^1 \| \nabla f[x^{k,i} + t(x^{k,i+1} - x^{k,i})] - \nabla f(x^{k,i}) \| \, \| x^{k,i+1} - x^{k,i} \| \, dt$$

$$f(x^{k,i}) - f(x^{k,i+1}) \geq (x_i^{k+1} - x_i^k)^2 \nabla_{ii} f(x^{k,i}) - \int_0^1 t L \| x^{k,i+1} - x^{k,i} \|^2 \, dt$$

(by (2.6) and the assumptions)

$$f(x^{k,i}) - f(x^{k,i+1}) \geq (x_i^{k+1} - x_i^k)^2 (\nabla_{ii} f(x^{k,i}) - L/2) \geq 0 . \tag{2.9}$$

ii) Since f is bounded from below, then from i) it follows that the sequence $\{f(x^{k,i})\}_k$ converges. Hence

$$\lim_{k \to \infty} (f(x^{k,i+1}) - f(x^{k,i})) = 0$$

and taking the limit of the expression (2.9)

$$0 = \lim_{k \to \infty} (f(x^{k,i}) - f(x^{k,i+1})) \geq \varepsilon \lim_{k \to \infty} (x_i^{k+1} - x_i^k)^2 .$$

It follows that

$$\lim_{k\to\infty} |x_i^{k+1}-x_i^k| = 0 \quad . \qquad \square$$

Corollary 2.2. Assume that H1, H2 and H3 hold.

$$\lim_{k\to\infty} (f(x^{k+1}) - f(x^k)) = 0$$

and

$$\lim_{k\to\infty} \|x^{k+1}-x^k\| = 0 \quad .$$

Proof. The result is an immediate consequence of the fact that

$$f(x^{k+1}) = f(x^{k+1,1}) < f(x^{k,n}) < f(x^{k,n-1}) < \ldots < f(x^{k,1}) = f(x^k) \quad . \square$$

The global convergence of the procedure is shown in the following theorem.

Theorem 2.3. Assume that H1, H2 and H3 hold. Any limit point $\bar{x}$ of the sequence $\{x^k\}$ generated by the procedure is a stationary point ; i.e. for all $x > 0$

$$\nabla f(\bar{x})^T(\bar{x}-x) \leq 0 \quad .$$

Proof. Let $\{x^{k_j}\}$ be a subsequence such that

$$\lim_{j\to\infty} x^{k_j} = \bar{x} \quad .$$

Denote $T_i : R^n \to R^n$ the mapping defined by

$$T_i(x^{k,i}) = x^{k,i+1}$$

where $x^{k,i+1}$ is specified in relation (2.3). Obviously, T_i $(1<i<n)$ are continuous mapping and hence the mapping T

$$T(x^k) = T_n \circ T_{n-1} \circ \ldots \circ T_1(x^k) = x^{k+1}$$

is also continuous. It follows that

$$\lim_{j\to\infty} T(x^{k_j}) = T(\bar{x})$$

and from Corollary 2.2

$$\|T(\bar{x})-\bar{x}\| = \lim_{j\to\infty} \|T(x^{k_j}) - x^{k_j}\| = 0 \quad .$$

Hence $T(\bar{x}) = \bar{x}$ and $T_i(\bar{x}) = \bar{x}$ for $1\leq i\leq n$. Furthermore, referring to relation (2.5), it follows that for $1\leq i\leq n$

$$\nabla_i f(\bar{x})\ (\bar{x}_i-y_i) \leq (\bar{x}_i-\bar{x}_i)\ \nabla_{ii} f(\bar{x})(\bar{x}_i-y_i) = 0$$

for all $y_i > 0$. Therefore, for all $y > 0$

$$\nabla f(\bar{x})^T(\bar{x}-y) = \sum_{i=1}^{n} \nabla_i f(\bar{x})(\bar{x}_i-y_i) \leq 0 \quad . \qquad \square$$

If the Lipschitz condition H2 is not satisfied, then the procedure can be modified slightly to use an Armijo type stepsize. Indeed, rather than using (2.3) to specify x_i^{k+1}, let

$$\tilde{x}_i^{k+1} = \text{Max}\left\{0, x_i^k - \frac{\nabla_i f(x^{k,i})}{\nabla_{ii} f(x^{k,i})}\right\}$$

and

$$x_i^{k+1} = x_i^k + \alpha(\tilde{x}_i^{k+1}-x_i^k) \quad .$$

It is easy to verify that this procedure converges globally if α is an Armijo type stepsize (see Zubieta, 1986).

This algorithm can be specialized to the spatial equilibrium problem. Denote by t the vector with components t_{ij}

$$d_i = \sum_{j=1}^{n} t_{ij} \qquad , 1\leq i\leq m$$

$$b_j = \sum_{i=1}^{m} t_{ij} \qquad , 1\leq j\leq n \quad .$$

At the k^{th} iteration of the procedure, denote by $t^{k,\mu\nu}$, $d_i^{k\nu}$ and $b_j^{k\mu}$ the vector t and the scalar d_i and b_j, respectively, evaluated at the current value of the variables t_{ij} when the variable $t_{\mu\nu}$ is to be modified. Denote

by S_i', I_j' and c_{ij}' the derivative of S_i, I_j and c_{ij}, respectively. The algorithm is then summarized as follows.

O-D ALGORITHM

Step 0. Let $t^0 > 0$ and $k = 0$.

Step 1. For all pairs (i,j),

$$t_{ij}^{k+1} = \text{Max}\left\{0, t_{ij}^k - \frac{S_i(d_i^{kj}) + c_{ij}(t_{ij}^k) - I_j(b_j^{ki})}{S_i'(d_i^{kj}) + c_{ij}'(t_{ij}^k) - I_j'(b_j^{ki})}\right\} \qquad (2.10)$$

Step 2. If $t^{k+1} = t^k$, stop. Otherwise, let $k = k + 1$ and repeat Step 1.

3. BLOCK DECOMPOSITION

In numerical analysis the Gauss-Seidel method for solving a system of equations can be extended to a block version where several variables are modified simultaneously (see Varga, 1962). The following extension is in this spirit.

Referring to problem (2.1), let $x = [x_1^T, x_2^T, \ldots, x_m^T]^T$ where $x_j \in R^{n_j}$, $n_j \geq 1$. A straightforward extension of the O-D pairs decomposition procedure is to take x_j^{k+1} as the optimal solution of the subproblem

$$\underset{x_j \geq 0}{\text{Min}} \; F_j(x^{k,j})^T(x_j - x_j^k) + \tfrac{1}{2}(x_j - x_j^k)^T H_j(x^{k,j})(x_j - x_j^k) \qquad (3.1)$$

where $F_j(x^{k,j})$ and $H_j(x^{k,j})$ are the gradient and the Hessian matrix of f at $x^{k,j}$ with respect to the components of x_j, respectively. Clearly, the efficiency of the procedure depends on the solution procedure chosen to solve (3.1).

For the spatial equilibrium problem there are two natural ways to form the blocks of variables :

i) (according to origin)

$$t = [t_{1.}^T, t_{2.}^T, \ldots, t_{m.}^T]^T$$

where $t_{i.} = [t_{i1}, t_{i2}, \ldots, t_{in}]^T$ is the vector of flows along the arcs out of origin i;

ii) (according to destination)

$$t = [t_{.1}^T, t_{.2}^T, \ldots, t_{.n}^T]^T$$

where $t_{.j} = [t_{1j}, t_{2j}, \ldots, t_{mj}]^T$ is the vector of flows along the arcs going in destination j.

Since subproblems (3.1) are not generally trivial to solve unless block dimension is one, we introduce two other procedures that can be analyzed in the context of active sets methods (see Luenberger, 1984).

3.1 Active sets approach

The basic idea in any active sets method is to solve a sequence of problems specified by different subsets of constraints required to be active. At each main iteration the subset of constraints required to be active is modified appropriately to generate a sequence of feasible solutions converging to an optimal solution for the original problem.

The procedure is summarized in the following algorithm.

ACTIVE SETS ALGORITHM

Step 0. Let $t^0 \geq 0$ be an initial solution for (1.7). Denote by

$$W = \{(i,j) : t_{ij}^0 = 0\}$$

the set of indices of the initial active constraints.

Step 1. Determine an optimal solution t^W for the problem

$$\text{Min } g(t)$$

$$\text{Subject to} \quad t_{ij} \geq 0 \ (i,j) \notin \bar{W} \qquad (3.2)$$

$$t_{ij} = 0 \ (i,j) \in \bar{W} \quad .$$

Denote by λ_{ij} the optimal multiplier associated with the constraint $t_{ij} = 0$ or $t_{ij} \geq 0$.

Step 2. If $\lambda_{ij} \geq 0$ for all $(i,j) \in \bar{W}$, then t^W is an optimal solution of (1.7), and stop. Otherwise, generate another starting solution $t^0 \geq 0$, a new set W (see below for details) and repeat Step 1. □

To complete Step 1 of the algorithm, another iterative procedure is used. It starts from t^0 and moves to a new point t' on the surface generated by the active constraints W and where

$$g(t') < g(t^0) \quad .$$

Furthermore, if $t'_{ij} = 0$ for some indices $(i,j) \notin W$. then these are introduced into W and the process is repeated from t' with the enlarged set W. Hence, at the end of the procedure, the original set W has been modified to become $\bar{W}$ (with $W \subset \bar{W}$), and t^W is an optimal solution for problem (3.2).

The specific procedure to determine t^W is now summarized in the following t-algorithm. At the k^{th} iteration, let $d_i^k = \sum_{j=1}^{n} t_{ij}^k$ and $b_j^{ki} = \sum_{\ell=1}^{i-1} t_{\ell j}^{k+1} + \sum_{\ell=i}^{m} t_{\ell j}^k$.

t - ALGORITHM

Step 0. Let $t^0 \geq 0$ and $k = 0$.

Step 1. For $i=1,2,\ldots,m$

i) $\Gamma_i = \{j : t_{ij}^k > 0\}$

ii)
$$h_i^k = \frac{-\sum_{j\in\Gamma_i} \frac{S_i(d_i^k)+c_{ij}(t_{ij}^k)-I_j(b_j^{ki})}{c'_{ij}(t_{ij}^k)-I'_j(b_j^{ki})}}{1 + S'_i(d_i^k) \sum_{j\in\Gamma_i} \frac{1}{c'_{ij}(t_{ij}^k)-I'_j(b_j^{ki})}} \qquad (3.3)$$

iii) for $j\in\Gamma_i$

$$y_{ij}^k = t_{ij}^k - \frac{S_i(d_i^k)+c_{ij}(t_{ij}^k)-I_j(b_j^{ki})+S'_i(d_i^k).h_i^k}{c'_{ij}(t_{ij}^k)-I'_j(b_j^{ki})} \qquad (3.4)$$

iv)
$$\beta_i^k = \begin{cases} 1 & \text{if } y_{ij}^k > t_{ij}^k \text{ for all } j\in\Gamma_i \\ \underset{j\in\Gamma_i}{\text{Min}} \left\{ \frac{t_{ij}^k}{t_{ij}^k - y_{ij}^k} : t_{ij}^k - y_{ij}^k > 0 \right\}, & \text{otherwise} \end{cases} \qquad (3.5)$$

v)
$$t_{ij}^{k+1} = \begin{cases} t_{ij}^k + \beta_i^k (y_{ij}^k - t_{ij}^k) & \text{if } j\in\Gamma_i \\ 0 & \text{if } j\in\Gamma_i \end{cases} \qquad (3.6)$$

Step 2. If $t^{k+1} = t^k$, stop. Otherwise $k = k + 1$ and repeat Step 1. □

It is easy to verify that the vector y_i^k with components y_{ij}^k, $j\in\Gamma_i$ is an optimal solution for the following unconstrained problem

$$\text{Min } G\Gamma_i(t^{k,i\cdot})^T(y_i - t_i^k) + \tfrac{1}{2}(y_i - t_i^k)^T H\Gamma_i(t^{k,i\cdot})(y_i - t_i^k)$$

where t_i^k is the vector with components t_{ij}^k, $j\in\Gamma_i$, and $G\Gamma_i$ and $H\Gamma_i$ are the gradient and the Hessian matrix of g with respect to the variable t_{ij}, $j\in\Gamma_i$, respectively.

Assume there exists a scalar $L > 0$ such that for all t and t',

$$\|\nabla g(t) - \nabla g(t')\| \leq L \|t - t'\| ,$$

and a scalar $\varepsilon > 0$ such that for all t

$$L/2 + \varepsilon < \mu(t)$$

where $\mu(t)$ is the smallest eigenvalue of $H\Gamma_i(t)$. Then using an argument

similar to the proof of Lemma 2.1 i) it is easy to verify that $g(t^k(y_i^k)) \leq g(t^{k,i\cdot})$, where $t^k(y_i^k)$ is obtained from $t^{k,i\cdot}$ by replacing t_{ij}^k by y_{ij}^k for all $j\in\Gamma_i$. Hence, since $0<\beta_i^k\leq 1$, it follows that

$$g(t^{k,i+1\cdot}) \leq g(t^{k,i\cdot})$$

with the convention that $t^{k,(i+1)\cdot} = t^{k+1,1\cdot}$ if $i=m$. Note finally that if the preceding Lipschitz condition is not verified, a stepsize of Armijo type can be used to generate y_i^k such that $g(t^k(y_i^k)) \leq g(t^{k,i\cdot})$.

To analyze the global convergence of the Active Sets Algorithm, it is important to notice that as long as for some index i there exists at least one index $j\in\Gamma_i$ such that $y_{ij}^k < t_{ij}^k$, then at least one more constraint becomes active in block i at the current point. Hence after a finite number of iterations K_0 the sets Γ_i remain constant for all $1\leq i\leq m$, or all variables become equal to 0 (i.e. $\Gamma_i = \emptyset$ for all $1\leq i\leq m$).

Assume that iteration K_0 is attained and that $\Gamma_i\neq\emptyset$ for at least one index $1\leq i\leq m$. Let $\Lambda = \{i:\Gamma_i=\emptyset,\ 1\leq i\leq m\}$, $\Theta = \{j:j\in\Gamma_i \text{ for some } i\in\Lambda\}$, and for $j\in\Theta$, $\Xi_j = \{i\in\Lambda:j\in\Gamma_i\}$. Also, let $\tilde{t}$ be the subvector of t composed of the elements t_{ij} for $i\in\Lambda$ and $j\in\Gamma_i$. Then the t -algorithm can be seen as a block Gauss-Seidel-Newton method to solve the system of nonlinear equations

$$S_i(d_i) + c_{ij}(t_{ij}) - I_j(b_j) = 0 \qquad j\in\Gamma_i,\ i\in\Lambda\ . \tag{3.7}$$

starting at $\tilde{t}^{K_0}$.

The underlying optimization problem associated with (3.7) is

$$\text{Min } G(\tilde{t}) = \sum_{i\in\Lambda}\int_0^{\sum_{j\in\Gamma_i} t_{ij}} S_i(x)dx + \sum_{i\in\Lambda}\sum_{j\in\Gamma_i}\int_0^{t_{ij}} c_{ij}(x)dx - \sum_{j\in\Theta}\int_0^{\sum_{i\in\Xi_j} t_{ij}} I_i(y)dy. \tag{3.8}$$

The global convergence of the block Gauss-Seidel-Newton method is analyzed in Schechter (1968). Sufficient conditions for the convergence of the method from a point $\tilde{t}^{K_0}$ are that

i) $\nabla^2 G(\tilde{t})$ is positive definite for all $\tilde{t}$;

ii) $\{\tilde{t} : G(\tilde{t}) < G(\tilde{t}^{K_0})\}$ is bounded.

Now, referring to problem (1.7), and the conditions imposed on $S_i(d_i)$, $I_j(b_j)$ and $c_{ij}(t_{ij})$, condition ii) above is satisfied if $c_{ij}(t_{ij})$ is bounded. But even if $\nabla^2 G(\tilde{t})$ is a principal submatrix of $\nabla^2 g(t)$, and even under the conditions imposed on $S_i(d_i)$, $I_j(b_j)$, i) is not necessarily verified. Hence i) is a stronger condition required for global convergence of the Active Sets Algorithm.

In Step 2 of the Active Sets Algorithm, if for all (i,j)

$$\lambda_{ij} = \nabla_{ij} g(t^W) = S_i(d_i^W) + c_{ij}(t_{ij}^W) - I_j(b_j^W) > 0$$

then the procedure stops and t^W is an optimal solution of (1.7).

Otherwise, if $\lambda_{ij} < 0$ for some $(i,j) \in \bar{W}$, then the objective function $g(t)$ can be reduced if some of these constraints are not taken as active. Hence a new initial solution $t^0 > 0$ is generated to repeat Step 1. This solution is generated as follows

$$t_{ij}^0 = \begin{cases} -\dfrac{S_i(d_i^W)+c_{ij}(t_{ij}^W)-I_j(b_j^W)}{S_i'(d_i^W)+c_{ij}'(t_{ij}^W)-I_j'(b_j^W)} & \text{if } (i,j)\in\bar{W} \text{ and } \lambda_{ij} < 0 \\ t_{ij}^W & \text{otherwise .} \end{cases}$$

In one sense, we are applying one iteration of the 0-D algorithm to the variables $(i,j) \in \bar{W}$ such that $\lambda_{ij} < 0$ (but we are not updating the vector t during the procedure to evaluate the functions in (2.10)). Moreover, if the Lipchitz condition H3 is not satisfied, we use a stepsize of Armijo type to define t_{ij}^0.

3.2 Early restarts

Instead of using the t-Algorithm of Section 3.1 to solve the original problem (1.7) on a surface generated by some active constraints, we complete only one iteration on each block and proceed immediately to modify the active set W. This strategy generates the following algorithm.

EARLY RESTARTS ALGORITHM

Step 0. Let $t^0 > 0$ be an initial solution for (1.7), and $k = 0$.

Step 1. For $i = 1, 2, \ldots, m$.

i) $\Gamma_i = \{j : t_{ij}^k > 0 \text{ or } [S_i(d_i^k)+c_{ij}(t_{ij}^k)-I_j(b_j^{ki})] < 0\}$.

ii) Determine h_i^k, y_{ij}^k $(j\varepsilon\Gamma_i)$, β_i^k and t_{ij}^{k+1} using relations (3.3), (3.4), (3.5) and (3.6), respectively, where Γ_i is defined in i).

Step 2. If $t^{k+1} = t^k$, STOP. Otherwise $k = k + 1$ and repeat Step 1. □

As argued by Luenberger (1984), the convergence of such a variant of the active set method is difficult (if not impossible) to analyze, but it may be efficient in practice.

Note that the sets Γ_i in Step 1 of the Early Restarts Algorithm are defined differently from those in Step 1 of the t-Algorithm. On the other hand, the underlying set W in Step 1 of the Early Restarts Algorithm corresponds to the same set in Step 1 of the Active Sets Algorithm.

Finally, a variant of the Early Restarts Algorithm is obtained by modifying ii) of Step 1 to determine the variables t_{ij} as follows

$$t_{ij}^{k+1} = \begin{cases} \text{Max } \{0, y_{ij}^k\} & \text{if } j\varepsilon\Gamma_i \\ 0 & \text{if } j\notin\Gamma_i \end{cases} . \qquad (3.9)$$

Hence β_i^k does not have to be evaluated. On the other hand, (3.9) implies that in fact a stepsize is taken in a direction different from $(y_i^k - t_i^k)$. This variant called the Heuristic Algorithm, is compared numerically with the other algorithms in the next section.

4. NUMERICAL RESULTS

The algorithms described in the preceding sections were implemented in FORTRAN IV on the CYBER 185 of the Computer Center of the Université de Montréal.

Twenty problems with the number of variables ranging from 25 to 400 were used to compare the algorithms numerically. In all these problems the supply functions have the form :

$$S_i(x) = \Pi_i + \left(\frac{d_0}{x}\right)^2$$

and the demand functions have the form :

$$I_j(x) = \frac{1}{\Theta_j} \log\left(\frac{2000}{x}\right).$$

Moreover, the transportation costs $c_{ij}(x_{ij})$ are all linear. These problems were also solved with MINOS (see Murtagh and Saunders, 1980) used as a benchmark.

In all the algorithms, the theoretical stopping criterion $t^{k+1} = t^k$ is replaced by the equilibrium conditions

$$|S_i(d_i^{k+1}) + c_{ij}(t_{ij}^{k+1}) - I_j(b_j^{k+1})| < \varepsilon \quad \text{if} \quad t_{ij}^{k+1} > 0$$

$$S_i(d_i^{k+1}) + c_{ij}(t_{ij}^{k+1}) - I_j(b_j^{k+1}) \geq 0 \quad \text{if} \quad t_{ij}^{k+1} = 0$$

where $d_i^{k+1} = \sum_{j=1}^{n} t_{ij}^{k+1}$ and $b_j^{k+1} = \sum_{j=1}^{m} t_{ij}^{k+1}$. The tolerance $\varepsilon = 10^{-5}$. Furthermore, the same initial solution

$$t_{ij}^0 = 10 \qquad \text{for all } (i,j)$$

is used for all algorithms. The results are summarized in Table 4.1.

The Active Sets and the Heuristic Algorithms are more efficient than the O-D Algorithm (except for problem 10.3) and MINOS. The performance of the first two algorithms are quite similar.

The O-D Algorithm is more efficient than MINOS for smaller problems (25 and 100 variables) but for larger problem (225 and 400) MINOS becomes more efficient.

Many of the test problems could not be solved with the Early Restarts Algorithm. Furthermore, for problems that it can solve it is not as efficient as the Heuristic Algorithm.

To summarize, on these test problems, the Active Sets and the Heuristic Algorithms were the most efficient. Moreover, their ease of implementation makes them very attractive procedures for solving spatial equilibrium problems.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to Michael Florian for his helpful collaboration.

This research was supported by the Natural Sciences and Engineering Research Council of Canada (N.S.E.R.C.) (grant A8312) and by the F.C.A.R. program of the ministère de l'Éducation du Québec (grant 0899).

Table 4.1

Problems		O-D Algorithm		Active Sets Algorithm		Early Restarts Algorithm		Heuristic Algorithm		MINOS
Number	Number of variables	Iter.[1]	Time[2]	Iter	Time	Iter	Time	Iter	Time	Time
5.1	25	273	1.604	29	0.240	86	0.438	53	0.259	1.481
5.2	25	66	0.387	15	0.105	39	0.220	32	0.154	1.341
5.3	25	368	2.070	65	0.321	67	0.316	48	0.220	1.441
5.4	25	158	0.870	13	0.098	111	0.457	83	0.334	1.204
5.5	25	65	0.369	33	0.180	37	0.185	24	0.123	1.281
10.1	100	108	2.141	34	0.571			28	0.351	3.789
10.2	100	185	3.638	76	1.020	DNR		68	0.761	3.893
10.3	100	77	1.476	289	3.160	139	1.750	83	0.953	3.555
10.4	100	121	2.391	77	1.020			25	0.326	3.591
10.5	100	360	7.268	252	2.920	DNR		39	0.479	4.105
15.1	225	466	19.925	49	1.600			189	4.117	11.234
15.2	225	86	3.678	56	1.700			84	1.925	14.163
15.3	225	405	17.363	95	2.400	DNR		113	2.544	11.812
15.4	225	557	23.440	292	6.300			DNR		12.369
15.5	225	200	8.549	135	3.100			47	1.142	10.554
20.1	400	601	43.929	271	7.900			152	5.608	26.575
20.2	400	[3] 700	50.700	79	3.800			368	13.319	22.700
20.3	400	[3] 700	50.519	229	8.200	DNR		110	4.100	22.121
20.4	400	381	27.683	78	3.800			216	7.778	23.341
20.5	400	442	32.124	378	12.600			141	5.150	23.517

[1] Iter = number of iterations

[2] Time = CPU in sec.

[3] $\varepsilon = 10^{-2}$

DNR = did not run

REFERENCES

ASMUTH, R., B.C. EAVES and E.L. PETERSON. 1979. Computing Economic Equilibria on Affine Networks with Lemke's Algorithm. Mathematics of Operations Research 4, 209-214.

FLORIAN, M. and M. LOS. 1982. A New Look at Static Spatial Price Equilibrium Models. Regional Science and Urban Economics 12, 579-597.

FRIESZ, T.L., R.L. TOBIN, and P.T. HARKER. 1981. Variational Inequalities and Convergence of Diagonalization Methods for Derived Demand Network Equilibrium Problems. Report CUE-FNEM-1981-10-1, Department of Civil Engineering, University of Pennsylvania, Philadelphia, PA.

FRIESZ, T.L., R.L. TOBIN, T.E. SMITH and P.T. HARKER. 1983. A Nonlinear Complementarity Formulation and Solution Procedure for the General Derived Demand Network Equilibrium Problem. Journal of Regional Science 23, 337-359.

JONES, R., R. SAIGAL and M.H. SCHNEIDER. 1984. A Variable Dimension Homotopy for Computing Spatial Equilibria. Operations Reseach Letters 3, 19-24.

JONES, R., R. SAIGAL and M. SCHNEIDER. 1985. Computing Nonlinear Network Equilibria", Mathematical Programming 31, 57-66.

LUENBERGER, D.G. 1984. Linear and Nonlinear Programming. Addison-Wesley, Reading, MA.

MacKINNON, J.G. 1975. An Algorithm for the Generalized Transportation Problem, Regional Science and Urban Economics 5, 445-464.

MATHIESEN, L. 1974. A Linear Complementarity Approach to General Equilibrium", Working Paper WP-SOL-74-9, Department of Operations Research, Stanford University, Stanford, CA.

MURTAGH, B. and M. SAUNDERS. 1978. Large-Scale Linearly Constrained Optimization. Mathematical Programming 14, 41-72.

MURTAGH, B. and M. SAUNDERS 1980. MINOS/AUGMENTED User's Manual. Technical Report SOL-80-1, Department of Operations Research, Stanford University, Stanford, CA.

NAGURNEY, A. 1985. Computational Comparisons of Spatial Price Equilibrium Methods. Working Paper. School of Management, University of Massachusetts, Amherst, MA.

ORTEGA, J.M. and W.C. RHEINBOLDT. 1970. Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York.

PANG, J.S. and P.S.C. LEE. 1981. A Parametric Linear Complementarity Technique for the Computation of Equilibrium Prices in a Single Commodity Spatial Model. Mathematical Programming 20, 81-102.

PANG, J.S. 1983. Solution of the General Multi-Commodity Spatial Equilibrium Problem by Variational and Complementarity Methods". Journal of Regional Science 24, 403-414.

POLITO, J., B.A. McCARL and T.L. MORIN. 1980. Solution of Spatial Equilibrium Problems with Bender's Decomposition. Management science 26, 593-605.

SAMUELSON, P.A. 1952. Spatial Price Equilibrium and Linear Programming. American Economic Review 42, 283-303.

SCHECHTER, S. 1968. Relaxation Methods for Convex Problems. Siam Journal of Numerical Analysis 5, 601-612.

TAKAYAMA, T. and G.G. JUDGE. 1964. Equilibrium among Spatially Separated Markets : A Reformulation. Econometrica 32, 510-524.

TAKAYAMA, T. and G.G. JUDGE. 1970. Alternative Spatial Equilibrium Models", Journal of Regional Science 10, 1-12.

TAKAYAMA, T. and G.G. JUDGE. 1971. Spatial and Temporal Price and Allocation Models. North-Holland, Amsterdam.

VARGA, R.S. 1962. Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, New Jersey.

ZUBIETA, L. 1986. Modèles d'équilibre spatial et méthodes de résolution. Ph. D. thesis, département d'informatique et de recherche opérationnelle, Université de Montréal, Montréal.